# CLAIMS

What is claimed is:

1. A method of identifying open reading frames (ORFs) in a genome of an organism comprising the steps of:

(A) collecting a genomic sequence of a first organism;

(B) comparing the genomic sequence of the first organism to one or more other genomic libraries comprising genomes of other organisms containing ORFs; and

(C) determining ORFs for the first organism based on the comparison.

2. The method of Claim 1, wherein the method uses a Basic Local Alignment Search Tool (BLAST) program.

3. The method of Claim 2, wherein the p-value for the BLAST program is less than 1.

4. The method of Claim 1, wherein the method uses a FASTA program or its equivalent.

5. The method of Claim 1, wherein the step of collecting genomic sequences excludes sequences comprising known ORFs of the first organism.

6. The method of Claim 1, wherein the first organism is a plant, a virus, a bacterium, a vertebrate, or an invertebrate.

7. The method of Claim 6, wherein the first organism is a vertebrate selected from the group consisting of primate, equine, bovine, caprine, ovine, porcine, feline, canine, lupine, camelid, cervidae, rodent, avian and ichthyes.

8. The method of Claim 7, wherein the primate is a human.

9. The method of claim 1, wherein the first organism is a fungi.

10. The method of Claim 9, wherein the first organism is a fungi selected from the group consisting of oomycota, chytridiomycota, zygomycota, ascomycota, basidiomycota and deuteromycota.

11. The method of Claim 10, wherein the ascomycota is *Saccharomyces* or *Schizosaccharomyces*.

12. The method of Claim 11, wherein the *Schizosaccharomyces* is *S. pombe*.

13. The method of Claim 11, wherein the *Saccharomyces* is *Saccharomyces cerevisiae*.

14. The method of Claim 1, wherein the smORF encodes a polypeptide less than 100 amino acids long.

15. The method of Claim 1, wherein the smORF encodes a polypeptide of 17 to 100 amino acids.

16. A method of identifying coding open reading frames (ORFs) of an organism comprising the steps of:

(A) collecting genomic sequences of a first organism;

(B) identifying stop-to-stop ORFs of the first organism;

(C) translating the stop-to-stop ORFs into polypeptide sequences;

(D) comparing the polypeptide sequences of the first organism to amino acid translations of genomic libraries comprising genomes of other organisms; and

(E) identifying, based on sequence identity, ORFs of the first organism that are present in the other organisms, wherein the identified ORFs are coding ORFs.

17. The method of Claim 16, wherein the method uses a BLAST program.

18. The method of Claim 17, wherein the BLAST program uses a p-value less than 1.

19. The method of Claim 16, wherein the method uses a FASTA program.

20. The method of Claim 16, wherein method excludes previously identified ORFs of the first organism.

21. The method of Claim 16, wherein the first organism is an eukaryote or a prokaryote.

22. The method of Claim 21, wherein the first organism is the eukaryote is a vertebrate selected from the group consisting of primate, equine, bovine, caprine, ovine, porcine, feline, canine, lupine, camelid, cervidae, rodent, avian, and ichthyes.

23. The method of Claim 22, wherein the primate is a human.

24. The method of Claim 16, wherein the first organism is a fungi.

25. The method of claim 24, wherein the first organism is a fungi selected from the group consisting of oomycota, chytridiomycota, zygomycota, ascomycota, basidiomycota and deuteromycotoa.

26. The method of claim 25, wherein the ascomycota is *Saccharomyces* or *Schizosaccharomyces*.

27. The method of Claim 26, wherein the *Schizosaccharomyces* is *S. pombe*.

28. The method of Claim 26, wherein the *Saccharomyces* is *Saccharomyces cerevisiae*.

29. A smORF selected from SEQ ID NOS: 1-119.

30. A smORF selected from the group of sequences consisting of smORF18 (SEQ ID NO: 4), smORF570 (SEQ ID NO: 96), smORF139 (SEQ ID NO: 36), smORF57 (SEQ ID NO: 13) or a biologically active fragment thereof, and optionally, a sequence required for an amplification reaction.

31. A smORF identified using the method of claim 1.

32. A vector comprising the smORF of Claim 31.

33. A cell comprising the vector of Claim 32.

34. A smORF encoding a polypeptide selected from the group consisting of SEQ ID NOS: 674-1345.

35. A smORF encoding a polypeptide of smORF18 (SEQ ID NO: 677), smORF57 (SEQ ID No: 776), smORF139 (SEQ ID NO: 799), or smORF570 (SEQ ID NO: 814).

36. An isolated polypeptide encoded by the smORF of Claim 31.

37. A nucleic acid that hybridizes to a sense or an antisense strand of the smORF of Claim 31.

38. An isolated polypeptide comprising SEQ ID NOS: 674-1345 or 1346.

39. The isolated polypeptide of Claim 36, wherein the polypeptide comprises SEQ ID NOS: 674-791 or 792.

40. An isolated polypeptide selected from the group consisting of smORF18 (SEQ ID NO: 677) and smORF 57 (SEQ ID NO: 776).

41. An antisense compound comprising 15 to 50 nucleobases, wherein at least 8 contiguous nucleobases are derived from a nucleic acid sequence selected from SEQ ID NO: 1-119.

42. The antisense compound of claim 41, wherein the at least 8 contiguous nucleobases are selected from smORF18 (SEQ ID NO: 4) and smORF57 (SEQ ID NO: 13).

43. The antisense compound of claim 41, wherein the antisense compound is an antisense oligonucleotide.

44. The antisense compound of claim 41, wherein the oligonucleotide comprises at least one modified internucleoside linkage.

45. The antisense compound of claim 41, wherein the oligonucleotide is a chimeric oligonucleotide.

46. The antisense compound of claim 43, wherein the antisense oligonucleotide comprises at least one modified nucleobase.

47. The antisense compound of claim 43, wherein the antisense oligonucleotide comprises a modified internucleoside linkage, a phosphorothioate linkage, a modified sugar moiety, or a modified nucleobase.

48. A method of inhibiting the expression of a smORF encoding a protein from Table 2 comprising administering an antisense compound which binds to a corresponding nucleic acid of Table 2.

49. A method of identifying an inhibitory compound to a protein encoded by the ORF identified by claim 1 comprising the steps of:

(a) contacting the protein encoded by the ORF or a biologically active fragment of the protein with a compound under conditions effective to promote specific binding between the protein and the compound; and

(b) determining whether the protein or biologically active fragment thereof bound to the compound; and

(c) determining whether the compound that binds to the protein further inhibits the activity of the protein.

50. The method of Claim 47, wherein the compound is a library selected from a group consisting of a combinatorial small organic library, a phage display library and a combinatorial peptide library.

51. A polypeptide or biologically active fragment thereof comprising at least 10 contiguous amino acids of SEQ ID NOS: 674-1346.

52. A composition comprising the polypeptide or biologically active fragment thereof of Claim 51 and a pharmaceutically acceptable carrier.

53. An antibody or immunologically active fragment thereof which recognizes and binds to a polypeptide or fragment of the polypeptide of Claim 51.

54. The antibody of Claim 53, wherein the antibody is a human antibody, a humanized antibody, a primatized antibody, a monoclonal antibody or a bispecific antibody.

55. The immunologically active fragment of the antibody of Claim 53, wherein the fragment is Fab, Fab', $F(ab')_2$, Fv, scFv, and Fd.

56. The antibody of Claim 53, wherein the antibody recognizes and binds to a polypeptide selected from the group consisting of SEQ ID NOS: 674-792.

57. The antibody of Claim 53, wherein the antibody binds to the protein of smORF18, smORF57, smOR139, smORF570.

58. A pharmaceutical composition comprising a nucleic acid of claim 29 and a pharmaceutically acceptable excipient.

59. A pharmaceutical composition comprising a polypeptide of claim 38 and a pharmaceutically acceptable excipient.